CLAIMS

What is claimed is:

5   1.   A method of valuing transformation between XML
documents comprising:
        a)   measuring data loss for a transformation operation
that transforms a source node in a source schema,
corresponding to a source XML document, to a target node in
10   a target schema, corresponding to a target XML document;
        b)   measuring a potential data loss for said
transformation operation;
        c)   scaling operands implemented in said transformation
operation to measure their impact on said data loss and
15   potential data loss; and
        d)   calculating a transformation cost considering said
data loss, potential data loss and impact of said operands
for said transformation operation.

20   2.   The method of valuing transformation as described
in Claim 1, further comprising:
        e)   generating a first sequence of transformation
operations to match and transform a plurality of source
nodes in a source tree of said source schema to a plurality
25   of target nodes in a target tree of said target schema.

        3.   The method of valuing transformation as described
in Claim 2, further comprising:
        f)   calculating a first transformation cost for said
30   sequence of transformation operations.

        4.   The method of valuing transformation as described
in Claim 2, further comprising:

f) generating a second sequence of transformation operations to match and transform said plurality of source nodes to said plurality of target nodes;

g) selecting said first or second sequence of

5   transformation operations that has a lower transformation cost.

5.   The method of valuing transformation as described in Claim 1, wherein a) comprises:

10   measuring preservation of source data associated with said source schema when transforming to said target schema.

6.   The method of valuing transformation as described in Claim 1, wherein d) further comprises:

15   automatically calculating said transformation cost.

7.   The method of valuing transformation as described in Claim 1, further comprising:

before a), modeling said source schema and said target

20   schema as a tree structure creating a source tree having a plurality of source nodes and a target tree having a plurality of target nodes.

8.   The method of valuing transformation as described

25   in Claim 1, wherein said source schema is a source document type definition (DTD) and said target schema is target DTD.

9.   The method of valuing transformation as described in Claim 2,

30   converting said sequence of transformation operations into an Extensible Stylesheet Language of Transformations (XSLT) script.

10.   A system for valuing transformation between XML
schemas comprising:

a data capacity measurement module for measuring a data
capacity (DC) gap value (DC(op);

5   a potential data capacity measurement module for
measuring a potential data capacity (PDC) gap value
(PDC(op);

an operand factor measurement module for measuring an
operand factor (Fac(op); and

10   a cost module for measuring a transformation cost for
measuring the preservation of data capacity contained within
a source XML document, that is associated with a source
schema, when transforming said source XML document to a
target XML document that is associated with a target schema
15   using a set of transformation operations.

11.   The system as described in Claim 10, wherein said
transformation cost measures the accuracy of transforming
said source schema to said target schema.

20

12.   The system as described in Claim 10, wherein said
(DC) gap value (DC(op)) measures the preservation of said
data capacity at a source node in said source schema when
matching and transforming said source node to a target node
25   in said target schema using said set of transformation
operations.

13.   The system as described in Claim 10, wherein said
potential data capacity (PDC) gap value (PDC(op)) measures
30   the potential that said source node that is transformed can
accommodate different data capacity than said source node.

14.   The system as described in Claim 10, wherein said operand factor (*Fac(op)*) measures the affect of operands on said DC and PDC gap values.

5   15.   The system as described in Claim 10, wherein said transformation cost relates said *DC* gap value, said *PDC* gap value, and said operand factor.

16.   The system as described in Claim 10, wherein said operand factor is related to a number, a size, and a property of operands associated with said set of transformation operations.

17.   The cost model as described in Claim 10, further comprising:

a total transformation cost comprising the summation of said transformation costs for matching and transforming each of a plurality of source nodes in a source tree of said source schema to a plurality of target nodes in a target tree of said target schema.

18.   A computer system comprising:
a processor: and
a computer readable memory coupled to said processor and containing program instructions that, when executed, implement a method of valuing transformation between XML documents comprising:

a)   measuring data loss for a transformation operation that transforms a source node in a source schema, corresponding to a source XML document, to a target node in a target schema, corresponding to a target XML document;

b)   measuring a potential data loss for said transformation operation;

48

c) scaling operands implemented in said transformation operation to measure their impact on said data loss and potential data loss; and

d) calculating a transformation cost considering said
5 data loss, potential data loss and impact of said operands for said transformation operation.

19. The method of valuing transformation as described in Claim 18, further comprising:
10 e) generating a first sequence of transformation operations to match and transform a plurality of source nodes in a source tree of said source schema to a plurality of target nodes in a target tree of said target schema.

15 20. The method of valuing transformation as described in Claim 19, further comprising:
f) calculating a first transformation cost for said sequence of transformation operations.

20 21. The method of valuing transformation as described in Claim 19, further comprising:
f) generating a second sequence of transformation operations to match and transform said plurality of source nodes to said plurality of target nodes;
25 g) selecting said first or second sequence of transformation operations that has a lower transformation cost.

22. The method of valuing transformation as described
30 in Claim 18, wherein a) comprises:
measuring preservation of source data associated with said source schema when transforming to said target schema.

23. The method of valuing transformation as described in Claim 18, wherein d) further comprises:

automatically calculating said transformation cost.

5   24. The method of valuing transformation as described in Claim 18, further comprising:

before a), modeling said source schema and said target schema as a tree structure creating a source tree having a plurality of source nodes and a target tree having a

10   plurality of target nodes.

25. The method of valuing transformation as described in Claim 18, wherein said source schema is a source document type definition (DTD) and said target schema is target DTD.

15

26. The method of valuing transformation as described in Claim 19,

converting said sequence of transformation operations into an Extensible Stylesheet Language of Transformations

20   (XSLT) script.